university of groningen

faculty of economics and business

16008-GEM

Reliability and Validity of the Happiness Approach to Measuring Preferences

André van Hoorn

SOM is the research institute of the Faculty of Economics & Business at the University of Groningen. SOM has six programmes:
- Economics, Econometrics and Finance
- Global Economics & Management
- Human Resource Management & Organizational Behaviour
- Innovation & Organization
- Marketing
- Operations Management & Operations Research

Research Institute SOM
Faculty of Economics & Business
University of Groningen

Visiting address:
Nettelbosje 2
9747 AE Groningen
The Netherlands

Postal address:
P.O. Box 800
9700 AV Groningen
The Netherlands

T +31 50 363 9090/3815

www.rug.nl/feb/research

university of groningen / faculty of economics and business

Reliability and Validity of the Happiness Approach to Measuring Preferences

André van Hoorn
University of Groningen, Faculty of Economics and Business, Department Global Economics and Management
a.a.j.van.hoorn@rug.nl

Reliability and Validity of the Happiness Approach to Measuring Preferences

André van Hoorn

University of Groningen, PO Box 800, 9700 AV, Groningen, the Netherlands.

Tel: +31 503 633 929, Fax: +31 503 632 341, E-mail: A.A.J.van.Hoorn@rug.nl.

Acknowledgements

Helpful comments by Sjoerd Beugelsdijk, Esther-Mirjam Sent and participants of the 2014 symposium Do I like what I prefer? Integrating Research on Attitudes and Preferences in Konstanz, Germany and the 2015 joint annual conference of the International Association for Research in Economic Psychology and the Society for Advancement of Behavioral Economics in Sibiu, Romania are gratefully acknowledged. This paper has developed out of joint work with Robbert Maseland. The usual disclaimer applies.

Reliability and Validity of the Happiness Approach to Measuring Preferences

Abstract

While the use of happiness data to estimate "utility" functions has some interesting advantages over stated and revealed preferences methods and is growing in popularity, evidence on the reliability and validity of the happiness approach to measuring preferences is lacking. Moving beyond the intuitive appeal of estimating happiness functions, I draw on the literature in psychology on so-called psychometric quality to examine the following two features of the happiness approach to measuring preferences: (i) do repeated samples and different measures of happiness or subjective well-being (SWB) render similar preferences (what is called reliability)?; and (ii) do SWB-based preference measures relate to other measures that capture similar constructs in a logical way (what is called construct validity)? Empirical evidence indicates that SWB-based preferences exhibit high intertemporal, test-retest stability and are highly consistent when measured using alternative indicators of SWB (reliability). Similarly, SWB-based preferences relate to stated and revealed preferences measures of similar constructs in expected ways (construct validity). Overall, I conclude that estimating happiness ("utility") functions provides a reliable and valid means for measuring people's preferences.

Keywords: Heterogeneous preferences; revealed preferences; stated preferences; attitudes; subjective well-being; happiness

1. Introduction

Building on advances in the study of happiness or, less colloquially, subjective well-being (SWB) (I use the terms interchangeably), economists are increasingly using data on self-reported happiness and life satisfaction to measure people's preferences (Alesina et al. 2004; Clark et al. 2005; Di Tella and MacCulloch 2005; Finkelstein et al. 2009, 2013; Van Hoorn and Maseland 2013). The idea is that SWB data can be used to estimate a happiness (or "utility") function, where the estimated coefficients for the determinants of happiness reflect people's preferences. The happiness approach, in turn, has several interesting advantages compared to more traditional revealed and stated preference approaches to measuring preferences. Chief among these advantages is the method's flexibility, particularly the possibility of eliciting preferences towards states of affairs that cannot be affected by individual choice or states of affairs that are not easily mimicked in a controlled laboratory setting (e.g., inflation, environmental pollution) (Di Tella and MacCulloch 2006, 2008; Graham 2010). In addition, the method involves minimum cognitive effort from respondents, as respondents do not need to think about any choices that they make in the laboratory or contemplate their answers to a complex battery of stated preferences survey items. Instead, respondents are only required to report on their own happiness levels; the connection with other (macro-level or individual-level) determinants of SWB is made ex-post by researchers seeking to study preferences concerning certain life domains or outcomes (Di Tella and MacCulloch 2006). SWB-based preferences have thus been used to shed novel light on such issues as partisan differences in preferences for inflation and unemployment (Di Tella and MacCulloch 2005) or the long-standing debate on a distinct Protestant work ethic (Van Hoorn and Maseland 2013). The method's advantages and the potential of insightful applications notwithstanding, there is no systematic assessment of the happiness approach's ability to render meaningful measures of differences in people's preferences. Although the idea of estimating "utility" functions using SWB or happiness data has much intuitive appeal, it remains unclear whether this method is really able to deliver.

To address this issue, I draw on psychology and its tradition of assessing the so-called psychometric quality of the large variety of latent mental constructs studied by psychologists, for instance, general

intelligence or personality. The psychometric quality of an indicator or of a specific measurement method or instrument comprises two features, reliability or precision and (construct) validity or accuracy (e.g., Thorndike 2004). I assess the reliability of the happiness approach by considering whether samples collected at different points in time and using different indicators of SWB render similar preferences. Similarly, I assess the validity of SWB-based preference measures by checking whether these measures relate to other measures—specifically stated preferences measures and self-reported behavior—that capture related constructs in a logical way. To make sure that my findings are relevant to the practice of estimating happiness functions and to avoid personal biases tainting any data collected, my empirical analysis draws on large-scale pre-existing data available from the European Social Survey (ESS) and, to a lesser extent, the European Values Study and World Values Survey (EVS-WVS). These two surveys are the source not only of data on different indicators of SWB needed to assess the reliability of SWB-based preference measures, but also for self-reports of behavior and stated preferences measures needed to assess the construct validity of SWB-based preference measures.

Results of a variety of empirical analyses provide strong evidence for the reliability and validity of the happiness approach to measuring preferences. Concerning reliability, results indicate that SWB-based preferences exhibit high intertemporal, test-retest stability and invariability with regard to the measure of SWB (self-reported happiness or life satisfaction) that one employs to estimate the happiness function, as the resulting preference measures are strongly correlated. Similarly, results show, among others, that individuals whose SWB is hurt more by unemployment are looking for a job more intensively than are individuals whose SWB is hurt less by unemployment, which indicates that SWB-based preference measures indeed measure the construct they are supposed to measure. Overall, I conclude that the happiness approach provides a reliable and valid means for researchers to measure differences in people's preferences.

Below, I first present some background information on the idea of psychometric quality or reliability and validity of measures of latent constructs and on the happiness approach to measuring preferences. Section 3 outlines my approach to assessing the reliability and validity of SWB-based preferences.

Section 4 presents details on the data that I use, both the indicators of SWB and the other measures used in the empirical analysis. Section 5 presents the results of the analyses of the reliability and validity of SWB-based preference measures, while Section 6 concludes.

Importantly, my concern in this paper is not with the question whether people's choices or consumption patterns tend to maximize their SWB (but see, for instance, Akay et al. 2015 and Benjamin et al. 2012 for insightful contributions to this matter). Similarly, my concern is not with the reliability and validity of indicators of SWB themselves, only with the use of such indicators in the measurement of people's preferences. To be complete, however, in Appendix A I also present a review of evidence on the reliability and validity of SWB indicators instead of SWB-based measures of differences in preferences.

2. Background

2.1. Reliability and validity as components of psychometric quality

Matching psychologists' tradition of measuring mental constructs through indirect methods, there is a whole field in psychology called psychometrics that is concerned with assessing the quality of measurement instruments and measured constructs. The field of psychometrics emphasizes two essential features or psychometric properties of a particular indicator or measurement instrument. These are the indicator's reliability—the indicator's consistency and its ability to give the same results in repeated measurement, and the indicator's (construct) validity—the extent to which the indicator is indeed able to capture the construct it purports to measure. Reliability and (construct) validity are alternatively referred to as precision, reproducibility, repeatability or variable error, and accuracy or constant error respectively. Reliability and validity are the standard quality requirements for psychometric constructs and are extensively dealt with in any textbook on measurement and methods in psychology (see, for example, Thorndike 2004). The classic references are Cronbach and Meehl (1955) and the handbook on *Standards for Educational and Psychological Testing* developed by the American Educational Research Association, the American Psychological Association, and the U.S. National Council on Measurement in Education (last revised in 2014). Together, the reliability and validity of a particular indicator reflect the

extent to which the indicator meaningfully captures an underlying trait or disposition. Applications of the idea of reliability and validity run widely. Psychological and educational tests, for instance, are extensively evaluated on criteria of reliability and validity. This holds for commonly used tests like those aiming to measure individuals' intelligence (IQ tests) or tests of school children's mathematical and verbal skills, including the American Scholastic Aptitude Test / Scholastic Assessment Test or SAT Reasoning Test (now SAT).

Theoretically, reliability and validity are clearly distinct concepts. The difference is often explained using a target analogy (Figure 1). Measurement is more valid or accurate the closer shots are to the actual target center, the bull's-eye. In addition, separate shots may well miss the target but, as a whole, they clearly identify the target. The left-hand side of Figure 1 illustrates this understanding of validity. Reliability or precision, on the other hand, refers to the variation in the various shots. The closer together different shots are, the higher precision is. Even if the shots miss the actual target, if they are grouped closely, precision is high. The right-hand side of Figure 1 illustrates this understanding of reliability. In the target analogy, reliability is typically seen as a prerequisite for validity: if individual shots are not close together, they cannot all be close to the bull's-eye either.

<Insert Figure 1 about here>

Reliability and validity can be checked in various ways. An oft-used approach by which to judge reliability is to look at so-called test-retest correlations. This test typically applies the same instrument to measure a certain construct to the same set of subjects but at different occasions. The idea is that reliability requires that separate measurements render the same variation in scores. Hence, if the correlation between two different measurements is high, we can conclude that the instrument or indicator is reliable. Another approach to assessing the reliability of an indicator is to look at its internal consistency. If multiple items or measured scores are used to capture a particular underlying construct, the interrelatedness of the separate items or scores can be used to say something about the reliability with

which this construct is measured. An indicator is only deemed reliable if the separate items are strongly correlated, evidencing high internal consistency. A specific test is to calculate Cronbach α (Cronbach 1951) for a set of items and compare the resulting score to standard criteria in the psychometric literature (e.g., DeVellis 2012; Kline 2000; Nunnally and Bernstein 1994).

Tests of validity, not unlike the concept of validity itself, come in diverse forms. For practical purposes, validity is often divided in different types of validity with accompanying tests. Two of the most prominent types are convergent validity and discriminant validity. Convergent validity refers to the correspondence between the indicator of interest and other indicators purporting to measure the same construct. If the measure of interest is to have convergent validity the results obtained should converge on the results obtained using these other measures. Discriminant validity considers how well the indicator correlates (or, usually, fails to correlate) with measures that capture something different.

In practice, it is not always possible to draw a clear distinction between reliability and validity. As a guideline, reliability checks tend to involve measuring the *same construct* using the *same method*. Validity checks on the other hand tend to involve either *different constructs* or *different methods* or both. Construct validity, for instance, is assessed by looking at the relationship between the original indicator and some other measure of the same construct but obtained using a different measurement method. For discriminant validity, the measured constructs differ and occasionally the method as well, although the test does not require use of different methods. Still, even with these guidelines in mind, the distinction between reliability and validity can be difficult to make. On the other hand, in many cases drawing a clear distinction is actually not particularly relevant, which is also the case for the analysis presented in this paper.

Finally, in strict terms, reliability and validity are properties of a particular measurement method or instrument and not of the indicators or scores obtained by using a particular instrument or a particular measurement approach. For my purpose, however, we can rely on the intuitive understanding of reliability and validity as essential properties either of SWB-based preferences or of the happiness approach to measuring preferences.

2.2. Measuring differences in preferences using happiness data

Economists' intuitive, non-formal understanding of preferences is as tastes, which involves mental entities including people's attitudes and goals but also the criteria by which they evaluate outcomes or different states of affairs (Hirschman 1984; Kahneman et al. 1999; Kahneman and Sugden 2005; Rubinstein 2006). In the context of choice, people choose the alternative they like the most, the alternative from which they expect to derive most utility. The intuitive understanding of preferences is not dependent on actual choice, however, as when a person likes Picasso's earlier work better than his later work, has an attitude towards past events or has a favorite color or animal. Following the intuitive understanding of preferences as tastes, measuring preferences is about finding out what people like and how much and what a particular outcome or state of affairs means to them.

The happiness approach to measuring preferences fits the intuitive understanding of preferences as tastes or the liking or valuing of something. Working backwards from an indicator of individuals' well-being, it is possible to say something about people's preferences. The only thing that a researchers needs to do is to relate measured well-being, specifically measured SWB, to factors affecting this well-being, which can be any number of things or states of affairs, individual factors such as income or employment status or country-level variables such as inequality or inflation (Di Tella and MacCulloch 2006). The estimated coefficients for the factors affecting SWB—large or small, positive or negative—show how much certain people like or dislike / value or disvalue something in terms of its contribution to their well-being. Moreover, estimating SWB or happiness functions for different groups allows one to measure differences in preferences between these groups, as reflected in differences in the estimated coefficients.

Theoretically, the happiness approach and the idea of SWB-based preferences are grounded in the work of Daniel Kahneman and collaborators on experienced utility (Dolan and Kahneman 2008; Kahneman 1999, 2000; Kahneman and Sugden 2005; Kahneman and Thaler 1991, 2006; Kahneman et al. 1997). Kahneman and co-authors distinguish between experienced utility—the "hedonic quality" of an outcome—and decision utility—the "weight of an outcome in a decision" (Kahneman et al. 1997: 375).

Indicators of SWB (e.g., self-reported happiness or life satisfaction) can be seen as measures of experienced utility (Alesina et al. 2004; Di Tella and MacCulloch 2005; Helliwell and Barrington-Leigh 2010; Rabin 1998; Van Hoorn et al. 2010) so that SWB data can be used to estimate the shape and content of "experienced utility functions."

As mentioned in the introduction, there is a growing literature applying the happiness approach to study people's preferences, particularly systematic differences in preferences between groups of people. Some of the preferences most studied by economists using SWB data—most often self-reported life satisfaction—concern the preferences for equality and for income. The different groups of people considered in these studies include different countries (notably Europe versus the U.S.), political left-wingers versus political right-wingers, and people in poor health versus people in good health.[1] Alesina et al. (2004) estimate heterogeneous happiness functions to shed light on differences in the size of the welfare state between Europe and the United States and on partisan preferences. They report that inequality has a larger negative effect on SWB in Europe than in the U.S. and that left-wingers are hurt more by inequality than right-wingers are but only in Europe. Partisan differences are also reported by Di Tella and MacCulloch (2005). They test models of partisan business cycles and find that the happiness of left-wing (right-wing) individuals is lowered more by unemployment (inflation) than by inflation (unemployment). Di Tella et al. (2010) report further left-right differences in the preference for social status and income. Lelkes (2006) and Finkelstein et al. (2009, 2013) are some other SWB-based studies of differences in the preference for income. The former reports that the effect of economic variables

[1] To be complete, SWB-based preferences have also been used in valuation exercises that seek to put a price tag on (environmental) externalities and other such public (dis)amenities lacking a market valuation (e.g., Van Praag and Baarsma 2005). Such exercises are not concerned with studying preferences per se, however, but with estimating the effect of a particular externality on individuals' SWB and comparing the size of this effect with the SWB effect of income as a way of calculating the cost of the externality in the absence of market demand information.

including income on happiness is smaller among the religious than among the non-religious. Finkelstein et al. (2009, 2013) use SWB-based preferences to examine the health-state dependence of the utility function. They find that the marginal utility of consumption, as measured by the effect of an individual's consumption on his or her happiness rating, declines with deteriorating health.

Other applications of the happiness approach have considered a Communist legacy in the preferences of (former) East and West Germans (Van Hoorn and Maseland 2010) and religious differences in the preference for having a job, specifically the existence of a distinct Protestant work ethic (Van Hoorn and Maseland 2013). Meanwhile, resembling the latter cross-country analysis, (cross-cultural) psychologists have combined SWB-based preferences with measures of differences in national culture, specifically individualism versus collectivism, in large-scale studies of cultural differences in, among others, the effect of marriage on individuals' happiness (Diener et al. 2000). Finally, it is good to note that the happiness approach is flexible enough to consider differences in preferences towards certain self-assessed mental states and not only preferences towards objectively observable outcomes. Diener and Diener (1995), for instance, use SWB data to study cultural differences in the preference for self-esteem, while Van Praag et al. (2003) consider group differences in the effect of satisfaction with certain life domains—for instance "satisfaction with work" or "satisfaction with household income"—on overall life satisfaction.

3. Assessing the psychometric quality of SWB-based preferences

Following the above discussion of psychometric quality, my assessment of the happiness approach's ability to render meaningful measures of preferences focuses on the reliability and validity of the preference measures obtained from estimating happiness functions for different groups of people. In this assessment, SWB-based preferences are reliable if differences in preferences between groups are consistent across different applications of the happiness approach. Similarly, SWB-based preferences are valid if they uncover the same kind of differences in preferences between selected groups as alternative

preferences measures do, specifically self-reports of behavior (revealed preferences) and stated-preference measures.

At the heart of my reliability and validity assessments is of course the estimating of a happiness function for different groups. In line with Alesina et al. (2004) and Van Hoorn and Maseland (2013), among others, for my assessment of the reliability of SWB-based preferences I consider differences in preferences between countries, which is a standard variable collected in many of the most popular surveys of SWB. For my assessment of the validity of SWB-based preferences, in contrast, I use data on the alternative preferences measures—available from the same surveys—to classify individuals as belonging to distinct groups. This way, I can check whether differences in the estimated happiness functions for these various groups coincide with the differences in preferences expected for these groups, given prior information on their self-reported behaviors and stated preferences. Moreover, it should be clear that the choice for the specific groups considered for my validity analysis is motivated by pragmatic considerations, namely the availability of survey items of self-reported behaviors and stated preferences covering the same domain. In similar fashion, while the set of preferences that can be studied using happiness data is broad, I have selected to focus my reliability and validity assessments on the preference for work. Again, the reason is pragmatic, as this particular preference domain has the matching measures of self-reported behaviors and stated preferences that I need to check whether SWB-based preference measures relate to other measures that capture related constructs in a logical way. At the same time, though, focusing on SWB-based preferences for work is far from unreasonable, given that the well-being effect of having a job versus being unemployed and differences in SWB-based preferences for work have been widely studied in the happiness literature (Clark and Oswald 1994; Hetschko et al. 2014; Van Hoorn and Maseland 2013; Winkelmann and Winkelmann 1998).

3.1. Approach to assessing the reliability of the happiness approach to measuring preferences

My approach to assessing the reliability of the happiness approach to measuring preferences involves two different ways of considering the consistency of preferences measured using SWB data. First, I

consider test-retest correlations of SWB-based preferences, meaning that I compare country differences in preferences measured at different times. If the happiness approach is reliable, SWB-based preferences measured at different times should result in a more or less similar picture of differences in preferences (of course assuming that preferences do not change radically over time). To make this assessment of the test-retest stability of SWB-based preferences, I examine a set of countries for which citizens have been asked about their well-being as well as several other features of their lives, notably their employment status, in different years. The first step of the assessment, then, is to divide the complete sample in an earlier subsample and in a later subsample. The second step is to estimate cross-country differences in the SWB effect of (not) having a job for each of these two subsamples. Finally, I calculate correlations between the country scores obtained for these two subsamples, where stronger correlations testify to higher test-retest stability and hence to higher reliability of the happiness approach to measuring preferences.

As a second way of assessing the reliability of the happiness approach to measuring preferences, I experiment with different indicators of SWB used in estimating the "utility" functions for the different countries in the analysis. Two common indicators of SWB are self-reported happiness and life satisfaction where especially the latter SWB indicator has been much used by economists to measure preferences (Clark et al. 2005; Di Tella and MacCulloch 2005). As mentioned in the appendix on measures of SWB, SWB comprises a cognitive and an affective component where life satisfaction is more associated with the cognitive component and happiness is more associated with the affective component (Diener et al. 1999). These differences between self-reported happiness and life satisfaction notwithstanding, I think that considering the correlation between preferences measured using these alternative SWB indicators helps shed light on the happiness approach's ability to render meaningful measures of people's preferences. The more consistent SWB-based preference measures are across different SWB indicators (what I label cross-indicator consistency), the more confidence we can have in these measures' reliability.

Meanwhile, a generic question regarding both these reliability assessments is how strong the intertemporal (test-retest) and cross-indicator correlations need to be before we can conclude that SWB-based preferences are indeed reliable. The psychometric literature does not provide a direct answer to this

question, only indirectly through established criteria for internal consistency / Cronbach α. My adaptation of these criteria is to calculate Cronbach α for the combination of the two different measures of SWB-based preferences that I obtain using either the earlier and the later subsample or the two alternative indicators of SWB. Theoretically, scores on Cronbach α can range between 0 and 1, where a score of 0 means that all variation in a particular measure is attributable to error while a score of 1 indicates that there is no variation in this measure that is due to error. Common interpretation of different values of Cronbach α are that α below 0.5 means unacceptably low internal consistency, that $\alpha \geq 0.5$ and $\alpha < 0.6$ means poor internal consistency, that $\alpha \geq 0.6$ and $\alpha < 0.7$ means questionable internal consistency, that $\alpha \geq 0.7$ and $\alpha < 0.8$ means respectable internal consistency, that $\alpha \geq 0.8$ and $\alpha < 0.9$ means very good internal consistency and that $\alpha \geq 0.9$ means excellent internal consistency (DeVellis 2012; Kline 2000; Nunnally and Bernstein 1994). Following these guidelines, I judge Cronbach α above 0.7 to mean that SWB-based preferences have sufficient test-retest stability or cross-indicator consistency to conclude that the happiness approach to measuring preferences is reliable. Still, as picking a threshold value for Cronbach α as a criterion remains subjective, in my results section I not only present Cronbach α's but also correlations between the different measures of country differences in the preference for work, thus facilitating readers to make their own assessment.

3.2. Approach to assessing the validity of the happiness approach to measuring preferences

My assessment of the validity of SWB-based preferences involves checking whether group differences in the preference for work coincide with group differences in the preference for work measured using either stated preferences or self-reported behavior (revealed preferences). The first step in this assessment is to use individuals' answers to a stated preference item concerning the importance of having a job or a self-reported behavior to distinguish two groups of individuals, individuals with a strong preference for work versus individuals with a weak preference for work. The second step is to estimate a happiness function for the two groups thus identified and the third step is to compare the estimated coefficients for these two groups. If SWB-based preferences are valid, the SWB effect of not having a job should differ

substantially between these two groups in a way that is consistent with the ex-ante distinction between having a strong versus a weak preference for work. Absent such differences, we may question the validity of SWB-based preferences (i.e., the ability of SWB-based preferences to capture the construct they are supposed to measure). The exact survey items that I use to make the distinction between groups required by my approach to assessing the validity of SWB-based preferences are presented in the next section. More information on these surveys and data sets is available from the surveys' respective websites, http://www.europeansocialsurvey.org and http://www.worldvaluessurvey.org.

4. Data and measures

4.1. Sources of data and samples

Data for my analysis come from the European Social Survey (ESS) and, to a lesser extent, the European Values Study and World Values Survey (EVS-WVS). Both these data sets are widely used in happiness research and include data on SWB, as well as various other characteristics of individuals, including whether they have a job or are currently unemployed. To assess the reliability of SWB-based preferences, I rely exclusively on data from the ESS. The reason is that the ESS has collected data for representative national samples for a variety of countries at regular, two-year intervals since 2002. My assessment of the validity of SWB-based preferences on the other hand uses data from both surveys, as the ESS includes a relevant item on job search behavior, while the EVS-WVS includes a relevant item asking respondents how important they think it is to have a job. The specific data sets that I use are the cumulative data files that cover Waves 1-6 of the ESS and Waves 1-5 of the combined EVS and WVS data sets respectively (European Social Survey 2014; European Values Study Group and World Values Survey Association 2006; World Values Survey Association 2009). However, depending on data availability, the sample of individuals that ends up being used in the analysis is typically smaller. For the assessment of the test-retest stability of SWB-based preferences, I limit the sample to countries that have been included in all six waves of the ESS, starting in 2002 and ending in 2014. As a result, my reliability

assessment covers between 16 and 32 countries. Similarly, the stated-preference measure needed for my validity assessment has only been included in Waves 3 and 5 of the EVS-WVS.

4.2. Measures

4.2.1. Subjective well-being

Both the ESS and the EVS-WVS measure SWB using standard questionnaire items. As an illustration, the ESS life satisfaction item reads: "All things considered, how satisfied are you with your life as a whole nowadays?," while the EVS-WVS life satisfaction items reads: "All things considered, how satisfied are you with your life as a whole these days?" Importantly, and as indicated, there is a great deal of evidence indicating that SWB measures such as those included in the ESS and EVS-WVS are able to render meaningful indicators (see Appendix A for a review).

4.2.2. Stated preference measures and self-reported behavior involving the importance of having a job

The stated-preference item that I use to measure the importance that different groups of people attach to having a job is provided by the EVS-WVS. Specifically, Waves 3 and 5 of the EVS-WVS include an item asking respondents whether they agree with the statement that "To fully develop your talents, you need to have a job." The answering scale for this item comprises five possible answers: 1 Strongly agree; 2 Agree; 3 Neither agree or disagree; 4 Disagree; 5 Strongly disagree. To create the two groups needed for my validity assessment, I take a group comprising all individuals that indicated that they agree strongly with this statement and a group of individuals that did not indicate agreeing strongly with this statement. If SWB-based preferences are valid, I expect that (not) having a job has a significantly stronger effect on SWB in the former group compared to the latter group.

The self-reported behavior that I use to measure the importance that different groups of people attach to having a job is provided by the ESS item asking individuals about their daily activities during the last seven days. This item asks individuals whether they were engaged in paid work, in education, retired, unemployed and looking for a job or unemployed and not looking for a job, among others. The latter two

answer categories are relevant to my analysis, as job search behavior can be seen to represent a revealed preference for having a job. Given that unemployed individuals that are looking for a job seem to have a stronger preference for work, I expect that for this group of individuals the negative SWB gap with individuals in paid work is larger compared to the size of this SWB gap for individuals that are not looking for a job. If SWB-based preferences are valid, unemployed individuals that are *looking for a job* should experience greater disutility or unhappiness from not having a job than individuals that are *not looking for a job* do.

4.2.3. Employment status

The final measure in my analysis of the reliability and validity of SWB-based preferences concerns having a job versus being unemployed as a factor affecting people's SWB. In the analyses that draw on data from the ESS, I construct this measure by using the item asking individuals about their daily activities described above. To create a dummy variable of employment status, I only consider individuals that are either in paid work or unemployed, meaning that I exclude individuals that are engaged in other activities such as education or retirement. Hence, the generic distinction that I make is between individuals with a job versus individuals without a job. On the other hand, as just discussed, for my validity assessment I create two distinct dummy variables, one dummy variable that distinguishes between employed individuals and unemployed individuals that are *not looking for a job*, and one dummy variable that distinguishes between employed individuals and unemployed individuals that are *looking for a job*.

For the analysis that involves EVS-WVS data, I draw on the EVS-WVS item asking individuals about their employment status. In this case, I only consider individuals that indicated that they are employed full-time versus individuals that indicated that they are unemployed, meaning that I exclude, among others, individuals that are self-employed, part-time employed or retired.

5. Results

5.1. Is the happiness approach reliable?

Table 1 present the results of the assessment of the reliability of SWB-based preferences, focusing on measured country differences in the preference for work. Results are presented in the form of a correlation table covering combinations of SWB-based preferences for work that are measured either using different indicators of SWB (satisfaction or happiness) or using SWB data collected at different intervals (2002-2006 versus 2008-2014). Hence, the results in Table 1 testify both to the test-retest stability of SWB-based preferences and to their consistency across alternative indicators of SWB (cross-indicator consistency).

<Insert Table 1 about here>

Results show that SWB-based preferences perform well both on cross-indicator consistency (Row 1, Column 1) and on intertemporal stability (test-retest correlation) (Rows 2-3, Columns 2-3). In all cases, Cronbach α is well above 0.7, my chosen threshold. In fact, Cronbach α's tend to be closer to 0.9 or higher. Correlations between the different preference measures are similarly strong. In comparison, test-retest correlations of 0.5 have been deemed sufficient for the purpose of measuring differences in cultural values between countries (Hofstede 2001: 53). Hence, the evidence convincingly demonstrates the happiness approach's ability to render reliable preference measures.

5.2. Is the happiness approach valid?

Based on the above evidence on the reliability of the happiness approach to measuring preferences, SWB-based preferences appear to be measuring something meaningful. Indeed, it seems unlikely to find such strong evidence of test-retest stability and cross-indicator consistency if the underling preference measure did not contain any information. Still, it is valuable to examine what exactly is measured by differences in the SWB effect of (not) having a job, which is the purpose of assessments of construct validity.

<Insert Tables 2a and 2b about here>

Tables 2a and 2b present the results of my validity assessment, comparing the SWB effect of unemployment across groups with different stated (Table 2a) and revealed preferences (Table 2b) for having a job. In general, we would expect that individuals that have claimed that having a job is important or whose self-reported behavior reveals a comparatively high dislike of being unemployed will experience significantly greater disutility from not having a job. And this is borne out by the results, which suggest that the disutility gap between groups with different stated and revealed preferences for work can be as high as 15% (0.199 / 1.27).

5.3. Discussion

Taken together, the above results provide strong evidence that the happiness approach is able to deliver, rendering meaningful measures of differences in people's preferences. However, a limitation of the above analysis is that the evidence concerns only the preference for work. Hence, a possible objection to this conclusion is that the reliability and validity of SWB-based preferences may not generalize to preferences concerning other domains. I agree that, in principle, the evidence presented until now cannot show the universal reliability and validity of SWB-based preferences. However, I also cannot come up with any theoretical rationale as to why my results would only hold for the domain of work but not for SWB-based preferences concerning other domains. Nevertheless, some supplementary evidence would strengthen our confidence in the happiness approach's generic ability to render meaningful measures of people's preferences.

Unfortunately, the available data only allow me to provide such evidence for the reliability of the happiness approach and not for the method's validity. The reason is that assessing the validity of SWB-based preferences concerning domains other than work requires matching stated preference measures and self-reported behaviors that are simply not available from the same kind of large-scale pre-existing

surveys as used for my main analyses. However, focusing on reliability only, it is possible to assess the happiness approach's generic ability to render meaningful measures of preferences also concerning other domains.

<Insert Tables 3a-3d about here>

Tables 3a-3d present results on the test-retest stability and cross-indicator consistency of SWB-based preferences for the following four outcomes / states of affairs: intensity of social activities, social status, work-life balance, and self-esteem. The reason for selecting these specific factors is that they have been widely studied in the SWB literature (e.g., Boyce et al. 2010; Di Tella et al. 2010; Diener and Diener 1995; Diener et al. 1999) and that they are available in multiple waves so as to allow measurement of country differences in SWB-based preferences at different points in time. Appendix B presents details on the measurement of these four variables and the survey items used, all of which have been included in two or more waves of the ESS. Bottom line is that Cronbach α's for SWB-based preferences concerning these other domains tend to be high as well, typically well above the 0.7 threshold. Hence, the results for these four selected factors strongly support the idea that SWB-based preferences are indeed universally reliable (and valid). In fact, Cronbach α's are frequently higher than those found for the preference for work (cf. Table 1), suggesting that my earlier results may have been conservative. Overall, there is broad evidence that SWB-based preferences contain systematic information and are able to measure meaningful differences between groups of people.

6. Concluding remarks

Does the happiness approach to measuring preferences render reliable and valid measures of preferences? Although economists are increasingly drawn to the idea of using data on subjective well-being (SWB) to estimate happiness or "utility" functions, measuring preferences this way only makes sense if the resulting measures are in fact meaningful. However, evidence on the reliability and validity of

the happiness approach to measuring preferences is lacking, the method's intuitive appeal notwithstanding. To fill this gap, this paper has drawn on standard methods for psychometric quality assessment in psychology to examine two key features of SWB-based preferences. The first feature, fitting the idea of reliability or precision, concerns the happiness approach's ability to render similar preferences across repeated samples and different SWB indicators. The second feature, fitting the idea of (construct) validity or accuracy, concerns the happiness approach's ability to render preference measures that relate to other measures that capture similar constructs, specifically stated preferences measures and self-reported behavior (revealed preferences), in a logical way. Empirical results indicate that the happiness approach to measuring preferences does well on both these counts. SWB-based preferences exhibit high intertemporal, test-retest stability and are highly consistent across different indicators of SWB used. Similarly, SWB-based preferences relate to stated and revealed preferences measures of similar constructs in expected ways. Hence, the conclusion that estimating happiness ("utility") functions provides a reliable and valid means for measuring people's preferences.

Still, though, there are features of the happiness approach to measuring preferences that are not assessed in this paper but that would warrant attention when using SWB data to measure preferences. Notably, the method of measuring preferences using SWB data would benefit quite a bit from having sufficiently large samples to estimate the happiness function. The reason is that having more observations enables more reliable estimation, which, in turn, will improve the psychometric quality of the preference measures obtained. Hence, even though the happiness approach is able, in principle, to render reliable and valid measures of preferences, there is no guarantee that it will always be able to do so in practice. My goal here is not to discourage application of the happiness approach to measuring preferences. However, I do think it valuable practice for studies of SWB-based preferences to consider, for instance, the consistency of empirical results across different SWB indicators as a matter of course.

Appendix A: Psychometric quality of measures of subjective well-being

As mentioned, this paper is not concerned with the psychometric quality of indicators of SWB but in the use of these indicators in the measurement of people's preferences. Nevertheless, to be complete, I use this appendix to present a brief review of the evidence on the reliability and validity of SWB indicators, adding a discussion of important related findings. Further surveys of assessments of the psychometric quality of SWB indicators can be found in Diener (1994), Diener et al. (1999), Frey and Stutzer (2002), Helliwell and Barrington-Leigh (2010), Kahneman and Krueger (2006), Nettle (2005), Layard (2005), and Van Hoorn (2008), among others.

Reliability of measures of subjective well-being

The reliability of SWB indicators has mostly been assessed by looking at test-retest correlations. This research finds that minor differences in circumstances and technical features of the specific questionnaire used can have a significant impact on reported levels of SWB (see Schwarz and Strack 1999: 62). Correspondingly, the test-retest correlation for single-item measures often does not exceed 0.60 when the same question is asked twice during a one-hour interview (ibidem; see also Andrews and Withey 1976). Kahneman and Krueger (2006: 7) report a test-retest correlation of 0.59 in a sample of 218 respondents interviewed and re-interviewed two weeks apart. Veenhoven (1996: 6) discusses the effect of the lag between times of asking on test-retest correlations. If respondents are asked twice in the same interview correlations are approximately 0.70; if the lag spans a week, test-retest correlation drops to about 0.60. He points out, however, that people seldom make large moves, e.g., from "satisfied" to "dissatisfied." Finally, a detailed study by Ehrhardt et al. (2000) examines panel data for Germany for the period 1984-1994. They find that at the start of the sample period the year-to-year correlation was 0.45 but that, for unknown reasons, it gradually increased to 0.54.

In general, the reliability of SWB measures as evidenced by test-retest correlations is lower than that found for common microeconomic variables such as personal income (Krueger and Schkade 2008). It appears to be more similar to the reliability of other subjective measures such as self-reported health (see, for example, Crossley and Kennedy 2002; Stewart et al. 1988; VanderZee et al. 1996). Next to the lag

between times of asking (for example one hour or two weeks), the specific measurement scale used also affects test-retest correlation, however. Studies show that more advanced measures such as multi-item scales produce more reliable SWB scores (see the overview in Krueger and Schkade 2008). Finally, the reliability of SWB indicators tends to vary with the specific component of SWB that is being measured. SWB comprises an affective component that is more closely associated with happiness and involves the presence/absence of positive and negative affect, and a cognitive component that is more closely associated with satisfaction and involves a conscious, information-based assessment of one's life (Diener et al. 1999). Indicators targeting the affective part of SWB appear to have lower test-retest correlations than indicators aiming to measure the cognitive part. Thus, measures of happiness tend be more susceptible to whimsical circumstances—the outcome of a soccer match and the weather, whether it is rainy or sunny, are famous examples (Schwarz and Strack 1999: 62)—than are measures of life satisfaction (Krueger and Schkade 2008).

Validity of measures of subjective well-being

The many available literature surveys (see above) testify to the attention SWB researchers have paid to the validity of their key empirical construct. The conclusion of the surveys is that indicators of SWB are quite valid. Table A.1 presents a brief overview of some commonly cited evidence on the validity of measures of SWB.

<Insert Table A.1 about here>

The apparent pervasiveness of research validating measures of SWB notwithstanding, the indicators have also received much criticism. Very often this criticism is based on the above-mentioned finding that whimsical (and relatively minor) circumstances like the outcome of a soccer match can have an impact on the level of SWB people report (Schwarz and Strack 1999: 62). The errors generally seem to be of a random rather than a structural nature, however, working both to increase and decrease SWB scores.

Consequently, the measuring instrument is unlikely to be systematically biased. Indeed, using large enough samples would go a long way in addressing possible problems introduced by contextual factors influencing the reported level of SWB (cf. Di Tella and MacCulloch 2006: 29; Krueger and Schkade 2008: 1843).

More on the meaning of subjective well-being

The findings from the SWB literature presented above support the conclusion that SWB measures are meaningful. There is less evidence about the meaning of these measures, other than that they are capturing something important, however. SWB remains largely an abstract concept. Some additional research goes some way in tackling this issue, however. Notably, by now, there is a substantial body of research that relates SWB to human physiology and brain activity. Levesque et al. (2003), for example, applied functional magnetic resonance imaging (fMRI) to analyze the neuroanatomical correlates of sad feelings in healthy children. Measured brain activity indicated that sad feelings are associated with significant bilateral activations of the midbrain, the medial prefrontal cortex, the anterior temporal pole and the right ventrolateral prefrontal. Ryff et al. (2004) present some preliminary findings indicating that people with higher SWB (specifically more meaning, purposeful engagement et cetera in their lives) have lower levels of daily salivary cortisol and pro-inflammatory cytokines. In addition, for these individuals the duration of REM sleep is longer than for individuals with lower levels of reported SWB. Urry et al. (2004) administered SWB questionnaires to people prior to analyzing their brain activity. A higher level of reported SWB was associated with greater left than right superior frontal activation. Steptoe et al. (2005) report that positive affect is associated with reduced neuroendocrine, inflammatory and cardiovascular activity. Positive affect also has a negative relationship to cortisol output during the day (controlling for other factors such as age and gender) and heart rate. During mental stress testing in the laboratory, people with higher positive affect had smaller plasma fibrinogen stress responses. Finally, a study by Rainville et al. (2006) finds that basic emotions are associated with distinctive patterns of cardio-respiratory activity.

This and other work on the physiological aspects of SWB tells us that SWB is not merely a construct that exists within people's minds. Rather, variations in SWB manifest themselves in observable biophysical phenomena as well as in people's self-reports. One can judge indicators of SWB meaningful also because they correlate with objective variables that have a solid biological foundation.

Appendix B: Measurement of factors affecting happiness used for supplementary analyses

<Insert Table A.2 about here>

References[2]

Akay, A., Bargain, O., Jara, H.X., 2015. Back To Bentham: Should We? Large-Scale Comparison of Decision versus Experienced Utility for Income-Leisure Preferences. ISER Working Paper Series, No. 2015-02.

Alesina, A., Di Tella, R., MacCulloch, R., 2004. Inequality and happiness: are Europeans and Americans different? Journal of Public Economics 88, 2009-2042.

* Andrews, F.M., Crandall, R., 1976. The validity of measures of self-reported well-being. Social Indicators Research 3, 1-19.

* Andrews, F.M., Withey, S.B., 1976. Social indicators of well-being. New York / London: Plenum Press.

Benjamin, D.J., Kimball, M.S., Heffetz, O., Rees-Jones, A., 2012. What do you think would make you happier? What do you think you would choose? American Economic Review 102, 2083-2110.

[2] References marked with an asterisk are only included in the review of the reliability and validity of indicators of SWB presented in Appendix A.

Boyce, C. J., Brown, G.D., Moore, S.C., 2010. Money and happiness rank of income, not income, affects life satisfaction. Psychological Science 21, 471-475.

* Campbell, A., Converse, P.E., Rodgers, W.L., 1976. The Quality of American Life. New York: Russell Sage Foundation.

Clark, A.E., Oswald, A.J., 1994. Unhappiness and unemployment. Economic Journal 104, 648-659.

Clark, A.E., Etilé, F., Postel-Vinay, F., Senik, C., Van der Straeten, K., 2005. Heterogeneity in Reported Well-Being: Evidence from Twelve European Countries. Economic Journal 115, C118-C132.

Cronbach, L.J., 1951. Coefficient alpha and the internal structure of tests. Psychometrika 16, 297-334.

Cronbach, L.J., Meehl, P.E., 1955. Construct Validity in Psychological Tests. Psychological Bulletin 52, 281-302.

DeVellis, R.F., 2012. Scale development: Theory and applications. Third edition. Sage: Los Angeles.

Di Tella, R., Haisken-De New, J., MacCulloch, R., 2010. Happiness adaptation to income and to status in an individual panel. Journal of Economic Behavior & Organization 76, 834-852.

Di Tella, R., MacCulloch, R., 2005. Partisan Social Happiness. Review of Economic Studies 72, 367-393.

Di Tella, R., MacCulloch, R., 2006. Some Uses of Happiness Data in Economics. Journal of Economic Perspectives 20(1), 25-46.

Di Tella, R., MacCulloch, R., 2008. Gross National Happiness as an Answer to the Easterlin Paradox. Journal of Development Economics 86, 22-42.

* Diener, E., 1994. Assessing Subjective Well-Being: Progress and Opportunities. Social Indicators Research 31, 103-157.

Diener, E., Diener, M., 1995. Cross-Cultural Correlates of Life Satisfaction and Self-Esteem. Journal of Personality and Social Psychology 68, 653-663.

Diener, E., Gohm, C., Suh, E., Oishi, S., 2000. Similarity of the Relations between Marital Status and Subjective Well-Being Across Cultures. Journal of Cross-Cultural Psychology 31, 419-436.

* Diener, E., Sandvik, E., Pavot, W., Gallagher, D., 1991. Response artifacts in the measurement of subjective well-being. Social Indicators Research 24, 35-56.

Diener, E., Suh, E.-M., Lucas, R.E., Smith, H.L., 1999. Subjective Well-Being: Three Decades of Progress. Psychological Bulletin 125, 276-302.

Dolan, P., Kahneman, D., 2008. Interpretations of Utility and Their Implications for the Valuation of Health. Economic Journal 118, 215-234.

* Ehrhardt, J.J., Saris, W.E., Veenhoven, R., 2000. Stability of life-satisfaction over time: Analysis of change in ranks in a national population. Journal of Happiness Studies 1, 177-205.

* Ekman, P., Davidson, R.J., Friesen, W.V., 1990. The Duchenne Smile: Emotional Expression and Brain Physiology II. Journal of Personality and Social Psychology 58, 342-353.

European Social Survey, 2014. Cumulative File, ESS 1-6. Data file edition 1.1. Norwegian Social Science Data Services, Norway - Data Archive and distributor of ESS data.

European Values Study Group, World Values Survey Association, 2006. European and World Values Surveys four-wave integrated data file, 1981-2004 v.20060423. Aggregate File Producer: ASEP/JDS, Madrid.

Finkelstein, A., Luttmer, E.F.P., Notowidigdo, M.J., 2009. Approaches to Estimating the Health State Dependence of the Utility Function. American Economic Review 99, 116-121.

Finkelstein, A., Luttmer, E.F.P., Notowidigdo, M.J., 2013. What Good Is Wealth Without Health? The Effect of Health on the Marginal Utility of Consumption. Journal of the European Economic Association 11, 221-258.

* Frey, B.S., Stutzer, A., 2002. Happiness and Economics. How the Economy and Institutions Affect Well-Being. Princeton / Oxford: Princeton University Press.

Graham, C., 2010. The Easterlin and Other Paradoxes: Why Both Sides of the Debate May Be Correct. In: Diener, E., Kahneman, D., Helliwell, J.F., Eds. International differences in well-being. Oxford: Oxford University Press.

Helliwell, J.F., Barrington-Leigh, C.P., 2010. Measuring and Understanding Subjective Well-Being. NBER Working Paper No. 15887. Cambridge, MA: National Bureau of Economic Research.

Hetschko, C., Knabe, A., Schob, R., 2014. Changing identity: Retiring from unemployment. Economic Journal 124: 149-166.

Hirschman, A.O., 1984. Against Parsimony: Three Easy Ways of Complicating Some Categories of Economic Discourse. American Economic Review 74, 89-96.

Hofstede, G., 2001. Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations. Second edition. Thousand Oaks, CA: Sage.

Kahneman, D., 1999. Objective Happiness. In: Kahneman, D., Diener, D., Schwarz, N., Eds. Well-Being: The Foundations of Hedonic Psychology. New York: Russell Sage Foundation.

Kahneman, D., 2000. Experienced utility and objective happiness: A moment-based approach. In: Kahneman, D., Tversky, A., Eds. Choices, Values and Frames. New York: Cambridge University Press and the Russell Sage Foundation.

* Kahneman, D., Krueger, A.B., 2006. Developments in the Measurement of Subjective Well-Being. Journal of Economic Perspectives 20(1), 3-24.

Kahneman, D., Ritov, I., Schkade, D., 1999. Economic preferences or attitude expressions? An analysis of dollar responses to public issues. Journal of Risk and Uncertainty 19, 220-242.

Kahneman, D., Sugden, R., 2005. Experienced Utility as a Standard of Policy Evaluation. Environmental and Resource Economics 32, 161-181.

Kahneman, D., Thaler, R.H., 1991. Economic Analysis and the Psychology of Utility: Applications to Compensation Policy. American Economic Review 81, 341-352.

Kahneman, D., Thaler, R.H., 2006. Anomalies: Utility Maximization and Experienced Utility. Journal of Economic Perspectives 20(1), 221-234.

Kahneman, D., Wakker, P., Sarin, R., 1997. Back to Bentham? Explorations of Experienced Utility. Quarterly Journal of Economics 112, 375-405.

* Kammann, R., Farry, M., Herbison, P., 1984. The analysis and measurement of happiness as a sense of well-being. Social Indicators Research 15, 91-115.

Kline, P., 2000. The handbook of psychological testing. Second edition. London: Routledge.

* Krueger, A.B., Schkade, D.A., 2008. The reliability of subjective well-being measures. Journal of Public Economics 92, 1833-1845.

* Larsen, R.J., Diener, E., Emmons, R.A., 1985. An evaluation of subjective well-being measures. Social Indicators Research 17, 1-18.

* Layard, R., 2005. Happiness: Lessons from a new science. London: Allen Lane / Penguin Books.

Lelkes, O., 2006. Tasting freedom: Happiness, religion and economic transition. Journal of Economic Behavior & Organization 59, 173-194.

* Levesque, J., Joanette, Y., Mensour, B., Beaudoin, G., Leroux, J.-M., Bourgouin, P., Beauregard, M., 2003. Neural correlates of sad feelings in healthy girls. Neuroscience 121, 545-551.

* Nettle, D., 2005. Happiness: The science behind your smile. Oxford / New York: Oxford University Press.

Nunnally, J., Bernstein, L., 1994. Psychometric theory. New York: McGraw-Hill.

* Pavot, W., Diener, E., 1993. The Affective and Cognitive Context of Self-Reported Measures of Subjective Well-Being. Social Indicators Research 28, 1-20.

Rabin, M., 1998. Psychology and Economics. Journal of Economic Literature 36, 11-46.

* Rainville, P., Bechara, A., Naqvi, N., Damasio, A.R., 2006. Basic emotions are associated with distinct patterns of cardiorespiratory activity. International Journal of Psychophysiology 61, 5-18.

* Rodgers, W.L., Regula Herzog, A., Andrews, F.M., 1988. Interviewing Older Adults: Validity of Self-Reports of Satisfaction. Psychology and Aging 3, 264-272.

Rubinstein, A., 2006. Lecture Notes in Microeconomic Theory. Princeton / Oxford: Princeton University Press.

* Schwarz, N., Strack, F., 1999. Reports of Subjective Well-Being: Judgmental Processes and Their Methodological Implications. In: Kahneman, D., Diener, D., Schwarz, N., Eds. Well-Being: The Foundations of Hedonic Psychology. New York: Russell Sage Foundation.

* Seidlitz, L., Diener, E., 1993. Memory for Positive Versus Negative Events: Theories for the Differences Between Happy and Unhappy Persons. Journal of Personality and Social Psychology 64, 654-664.

Thorndike, R.M., 2004. Measurement and Evaluation in Psychology and Education. Seventh Edition. Saddle River, NJ: Prentice Hall.

* Urry, H.L., Nitschke, J.B., Dolski, I., Jackson, D.C., Dalton, K.M., Mueller, C.J., Rosenkranz, M.A., Ryff, C.D., Singer, B.H., Davidson, R.J., 2004. Making a Life Worth Living: Neural Correlates of Well-Being. Psychological Science 15, 367-372.

Van Hoorn, A., 2008. A Short Introduction to Subjective Well-being: Measurement, Correlates and Policy Uses. In: OECD, Ed. Statistics, Knowledge and Policy 2007: Measuring and Fostering the Progress of Societies. Paris: OECD Publishing.

Van Hoorn, A., Mabsout, R., Sent, E.-M., 2010. Happiness and Capability: Introduction to the Symposium. Journal of Socio-Economics 39, 339-343.

Van Hoorn, A., Maseland, R., 2010. Cultural differences between East and West Germany after 1991: communist values versus economic performance? Journal of Economic Behavior & Organization 76, 791-804.

Van Hoorn, A., Maseland, R., 2013. Does A Protestant Work Ethic Exist? Evidence from the Well-Being Effect of Unemployment. Journal of Economic Behavior & Organization 91, 1-12.

Van Praag, B.M.S., Baarsma, B., 2005. Using Happiness Surveys to Value Intangibles: The Case of Airport Noise. Economic Journal 115, 224-246.

Van Praag, B.M.S, Frijters, P., Ferrer-i-Carbonell, A., 2003. The anatomy of subjective well-being. Journal of Economic Behavior & Organization 51, 29-49.

* VanderZee, K.I., Sanderman, R., Heyink, J., 1996. A comparison of two multidimensional measures of health status: The Nottingham Health Profile and the RAND 36-Item Health Survey 1.0. Quality of Life Research 5, 165-174.

* Veenhoven, R., 1996. Developments in satisfaction research. Social Indicators Research, 37 1-46.

* Watson, D., Clark, L.A., 1991. Self Versus Peer Ratings of Specific Emotional Traits: Evidence of Convergent and Discriminant Validity. Journal of Personality and Social Psychology 60, 927-940.

Winkelmann, L., Winkelmann, R., 1998. Why are the unemployed so unhappy? Evidence from panel data. Economica 65, 1-15.

World Values Survey Association. 2009. World Values Survey 1981-2008 Official Aggregate v.20090901. Aggregate File Producer: ASEP/JDS, Madrid.

Table 1
Consistency of differences in the preference for work.

		Consistency across indicators of subjective well-being	Intertemporal (test-retest) stability	
		Column 1 Country satisfaction effect of having a job	Column 2 Country satisfaction effect of having a job, 2002-2006	Column 3 Country happiness effect of having a job, 2002-2006
Row 1	Country happiness effect of having a job [n=32]	<u>α = 0.941</u> <u>r = 0.907</u>		
Row 2	Country satisfaction effect of having a job, 2008-2014 [n=16]		**α = 0.870** **r = 0.781**	
Row 3	Country happiness effect of having a job, 2008-2014 [n=16]			**α = 0.878** **r = 0.787**

Notes: Table reports Cronbach α's and correlation coefficients for differences in the SWB effect of work in different countries. The number of countries/observations is in square brackets. Main results concerning the test-retest stability and cross-indicator consistency of measured country differences in the preference for work are in **bold** and <u>underlined</u> respectively. The analysis of cross-indicator consistency (Row 1, Column 1) involves all 32 countries included in one of the six waves of the ESS (2002-2014): Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Lithuania, Luxembourg, Netherlands, Norway, Poland, Portugal, Russia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, and U.K. The 16 countries included in the analysis of test-retest stability (Rows 2-3, Columns 2-3) are all the countries that have been surveyed in each of the six waves of the ESS: Belgium, Denmark, Finland, France, Germany, Hungary, Ireland, Netherlands, Norway, Poland, Portugal, Slovenia, Spain, Sweden, Switzerland, and U.K. The country scores underlying the results in this table are obtained by modeling SWB as a function of having a job or not, separately for each country in the sample. Individual-level data used are from the ESS.

Table 2a
Differences in the disutility of unemployment between groups with different stated preferences for having a job.

	Group 1: Job **crucial** to develop talent	Group 2: Job **not needed** to develop talent	Disutility difference: **Group 1 - Group 2**
Unemployed (1=yes)	-1.27 (0.04)	-1.07 (0.03)	-0.199 [SE = 0.048]
Intercept	6.73 (0.02)	6.90 (0.01)	
No. of observations	20,532	41,366	

Notes: Table reports estimated coefficients with standard errors in parentheses. Group 1 comprises employed and unemployed individuals that all indicated to agree strongly with the statement that having a job is needed to develop one's talents. Group 2 comprises employed and unemployed individuals that did not agree strongly with the statement that having a job is needed to develop one's talents. Dependent variable in regression equation is life satisfaction (1-10). Data used are from the WVS, Waves 3 and 5.

Table 2b
Differences in the disutility of unemployment for individuals with different job search behavior.

	Group 1: Employed individuals & unemployed individuals **looking** for a job	Group 2: Employed individuals & unemployed individuals **not looking** for a job	Disutility difference: **Group 1 - Group 2**
Not having a job (1=yes)	-1.61 (0.02)	-1.39 (0.03)	-0.219 [SE = 0.038]
Intercept	6.99 (0.01)	6.99 (0.01)	
No. of observations	151,623	144,850	

Notes: Table reports estimated coefficients with standard errors in parentheses. Group 1 comprises individuals in paid work and individuals not in paid work but *looking for a job*. Group 2 comprises individuals in paid work and individuals not in paid work and *not looking for a job*. Dependent variable in regression equation is life satisfaction (0-10). Data used are from the ESS, Waves 1-6.

Table 3a
Consistency of country differences in the preference for social activities.

		Consistency across indicators of subjective well-being	Intertemporal (test-retest) stability	
		Column 1 Country satisfaction effect of social activities	Column 2 Country satisfaction effect of social activities, 2002-2006	Column 3 Country happiness effect of social activities, 2002-2006
Row 1	Country happiness effect of social activities [n=32]	<u>α = 0.951</u> <u>r = 0.910</u>		
Row 2	Country satisfaction effect of social activities, 2008-2014 [n=16]		**α = 0.883** **r = 0.795**	
Row 3	Country happiness effect of social activities, 2008-2014 [n=16]			**α = 0.934** **r = 0.878**

Notes: Table reports Cronbach α's and correlation coefficients for differences in the SWB effect of social activities in different countries. The number of countries/observations is in square brackets. Main results concerning the test-retest stability and cross-indicator consistency of measured country differences in the preference for work are in **bold** and <u>underlined</u> respectively. As in Table 1, the analysis of cross-indicator consistency (Row 1, Column 1) involves all 32 countries included in one of the six waves of the ESS (2002-2014), while the analysis of test-retest stability (Rows 2-3, Columns 2-3) involves the 16 countries that have been surveyed in each of the six waves of the ESS. The country scores underlying the results in this table are obtained by modeling SWB as a function of the frequency of engaging in social activities, separately for each country in the sample. Individual-level data used are from the ESS.

Table 3b
Consistency of country differences in the preference for social status.

		Consistency across indicators of subjective well-being	Intertemporal (test-retest) stability	
		Column 1 Country satisfaction effect of social status	Column 2 Country satisfaction effect of social status, 2002-2006	Column 3 Country happiness effect of social status, 2002-2006
Row 1	Country happiness effect of social status [n=32]	<u>α = 0.972</u> <u>r = 0.945</u>		
Row 2	Country satisfaction effect of social status, 2008-2014 [n=16]		**α = 0.938** **r = 0.884**	
Row 3	Country happiness effect of social status, 2008-2014 [n=16]			**α = 0.918** **r = 0.855**

Notes: Table reports Cronbach α's and correlation coefficients for differences in the SWB effect of social status in different countries. The number of countries/observations is in square brackets. Main results concerning the test-retest stability and cross-indicator consistency of measured country differences in the preference for work are in **bold** and <u>underlined</u> respectively. As in Table 1, the analysis of cross-indicator consistency (Row 1, Column 1) involves all 32 countries included in one of the six waves of the ESS (2002-2014), while the analysis of test-retest stability (Rows 2-3, Columns 2-3) involves the 16 countries that have been surveyed in each of the six waves of the ESS. The country scores underlying the results in this table are obtained by modeling SWB as a function of one's social status measured by income rank, separately for each country in the sample. Individual-level data used are from the ESS.

Table 3c

Consistency of country differences in the preference for work-life balance.

		Consistency across indicators of subjective well-being	Intertemporal (test-retest) stability	
		Column 1 Country satisfaction effect of satisfaction with work-life balance	Column 2 Country satisfaction effect of satisfaction with work-life balance, 2006	Column 3 Country happiness effect of satisfaction with work-life balance, 2006
Row 1	Country happiness effect of satisfaction with work-life balance [n=28]	<u>α = 0.939</u> <u>r = 0.890</u>		
Row 2	Country satisfaction effect of satisfaction with work-life balance, 2014 [n=22]		**α = 0.784** **r = 0.653**	
Row 3	Country happiness effect of satisfaction with work-life balance, 2014 [n=22]			**α = 0.798** **r = 0.683**

Notes: Table reports Cronbach α's and correlation coefficients for differences in the SWB effect of satisfaction with work-life balance in different countries. The number of countries/observations is in square brackets. Main results concerning the test-retest stability and cross-indicator consistency of measured country differences in the preference for work are in **bold** and <u>underlined</u> respectively. The analysis of cross-indicator consistency (Row 1, Column 1) involves the 30 countries included in either Wave 3 (2006) or Wave 6 (2014) of the ESS: Austria, Belgium, Bulgaria, Switzerland, Cyprus, Czech Republic, Germany, Denmark, Estonia, Spain, Finland, France, United Kingdom, Hungary, Ireland, Israel, Iceland, Italy, Lithuania, Netherlands, Norway, Poland, Portugal, Russia, Sweden, Slovenia, Slovakia and Ukraine. The 22 countries included in the analysis of test-retest stability (Rows 2-3, Columns 2-3) are all the countries that have been surveyed in both Wave 3 and Wave 6 of the ESS: Belgium, Bulgaria, Switzerland, Cyprus, Germany, Denmark, Estonia, Spain, Finland, France, Hungary, Ireland, Netherlands, Norway, Poland, Portugal, Russia, Sweden, Slovenia, Slovakia, Ukraine, and U.K. The country scores underlying the results in this table are obtained by modeling SWB as a function of satisfaction with the work-life balance offered by one's job, separately for each country in the sample. Individual-level data used are from the ESS.

Table 3d

Consistency of country differences in the preference for self-esteem.

		Consistency across indicators of subjective well-being	Intertemporal (test-retest) stability	
		Column 1 Country satisfaction effect of self-esteem	Column 2 Country satisfaction effect of self-esteem, 2006	Column 3 Country happiness effect of self-esteem, 2006
Row 1	Country happiness effect of self-esteem [n=28]	<u>α = 0.971</u> <u>r = 0.943</u>		
Row 2	Country satisfaction effect of self-esteem, 2014 [n=22]		**α = 0.776** **r = 0.692**	
Row 3	Country happiness effect of self-esteem, 2014 [n=22]			**α = 0.840** **r = 0.782**

Notes: Table reports Cronbach α's and correlation coefficients for differences in the SWB effect of self-esteem in different countries. The number of countries/observations is in square brackets. Main results concerning the test-retest stability and cross-indicator consistency of measured country differences in the preference for work are in **bold** and <u>underlined</u> respectively. As in Table 3c, the analysis of cross-indicator consistency (Row 1, Column 1) involves all 28 countries included in in either Wave 3 (2006) or Wave 6 (2014) of the ESS, while the analysis of test-retest stability (Rows 2-3, Columns 2-3) involves the 22 countries that have been surveyed in both Wave 3 and Wave 6 of the ESS. The country scores underlying the results in this table are obtained by modeling SWB as a function of self-esteem, separately for each country in the sample. Individual-level data used are from the ESS.

Table A.1
Validity of measures of subjective well-being.

Study	Main findings
Andrews and Crandall (1976)	Different single-item measures of SWB give similar results and of the total variance in single-item scales, about 64% is valid variance.[#] Validity can be increased further by using composite multi-item indicators; a five-item scale would typically have about 80% valid variance.
Costa and McCrae (1988)	Self-reported SWB scores show convergence with reports filed by spouses.
Diener et al. (1991)	Self-reported levels of SWB converge on those reported by others (e.g., by family and friends).
Ekman et al. (1990)	Subjective reports on the experience of positive emotions during an experiment correlated with the number of genuine smiles (referred to as "Duchenne" smiles) that occurred during the experiment.
Kammann et al. (1984)	Factor analysis shows that 13 different SWB scales measure the same general well-being factor. In addition, the negative halves of these scales measure the same construct as scales of neuroticism, depression and so-called trait anxiety.
Larsen et al. (1985)	The authors consider the validity of various SWB measures and report strong negative relations between positive overall SWB on the one hand and neuroticism and self-reported symptoms on the other; and strong positive correlations between overall SWB and ratings for several life domains, such as friendship, love life and financial situation.
Pavot and Diener (1993)	Self-reported SWB correlates strongly with peer reports, the personality traits of extraversion and neuroticism but not with current mood. Context has a slight effect on reported scores when single-item measures are used but multi-item scales appear largely immune.
Rodgers et al. (1988)	A comparison of SWB scores obtained using different indicators shows that more than 50% of total variance is valid variance and directly related to the construct of interest while only 10% is attributable to method (and one third is due to measurement error in the specific indicator used).
Seidlitz and Diener (1993)	For people with higher SWB scores, memories of happy events are more accessible than for people with lower SWB, while current mood has only a modest effect on the availability of memories.
Watson and Clark (1991)	There is a strong correlation between self-reported scores on eight different negative affect and positive affect scales and ratings by peers.

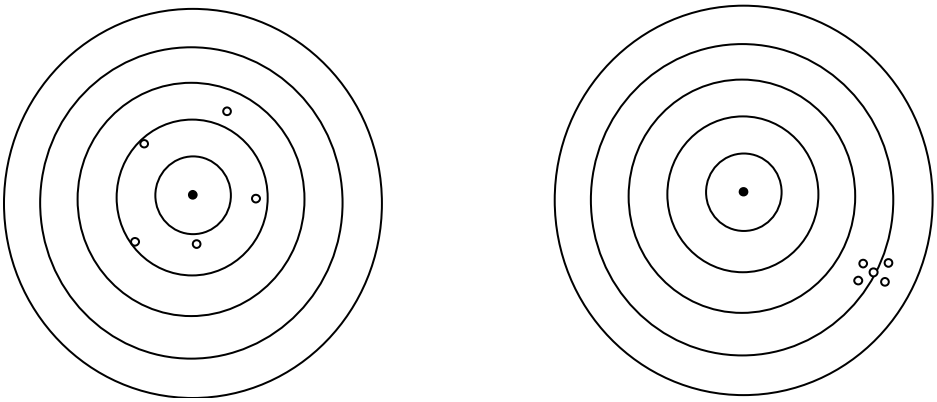
[#] The difference in SWB as actually perceived by the different respondents constitutes valid variance. Other (non-valid) variance results from measurement error, for instance, from respondents entertaining a different understanding of the response categories (being, say, "very happy" does not mean the same thing to everybody).

Table A.2
Measures of factors affecting happiness used for supplementary analyses.

Variable name	Description
Intensity of social activities	Measured using the item asking individuals: "Compared to other people of your age, how often would you say you take part in social activities?" Possible answers are: 1 Much less than most; 2 Less than most; 3 About the same; 4 More than most; and 5 Much more than most. Measure is available for ESS Waves 1-6.
Social status	Measured as rank of household income (cf. Boyce et al. 2010), which integrates the items measuring income included in different waves of the ESS. Waves 1-3 classified respondents' answers to the item on household income using a 12-point scale, while Waves 4-6 used a 10-point scale. The measure of rank income is created by calculating respondents' income percentile relative to respondents from the same country surveyed in the same year/wave Measure is available for ESS Waves 1-6.
Satisfaction with work-life balance	Measured using the item asking individuals: "How satisfied are you with the balance between the time you spend on your paid work and the time you spend on other aspects of your life?" The answer scale ranges from 0, Extremely dissatisfied to 10, Extremely satisfied. Measure is available for ESS Waves 3 and 6.[$]
Self-esteem	Measured using the item asking individuals how much they agree with the following statement: "In general I feel very positive about myself." Possible answers are: 1 Agree strongly; 2 Agree; 3 Neither agree nor disagree; 4 Disagree; and 5 Disagree strongly. Measure is available for ESS Waves 3 and 6.

Notes: All variables are treated as continuous measures in the analyses.
[$] Wave 5 of the ESS includes an almost identical item. However, to be consistent, I only use data collected in Waves 3 and 6.

Figure 1
Validity versus reliability in psychometric evaluation: The target analogy.



Validity or accuracy *Reliability or precision*

List of research reports

12001-HRM&OB: Veltrop, D.B., C.L.M. Hermes, T.J.B.M. Postma and J. de Haan, A Tale of Two Factions: Exploring the Relationship between Factional Faultlines and Conflict Management in Pension Fund Boards

12002-EEF: Angelini, V. and J.O. Mierau, Social and Economic Aspects of Childhood Health: Evidence from Western-Europe

12003-Other: Valkenhoef, G.H.M. van, T. Tervonen, E.O. de Brock and H. Hillege, Clinical trials information in drug development and regulation: existing systems and standards

12004-EEF: Toolsema, L.A. and M.A. Allers, Welfare financing: Grant allocation and efficiency

12005-EEF: Boonman, T.M., J.P.A.M. Jacobs and G.H. Kuper, The Global Financial Crisis and currency crises in Latin America

12006-EEF: Kuper, G.H. and E. Sterken, Participation and Performance at the London 2012 Olympics

12007-Other: Zhao, J., G.H.M. van Valkenhoef, E.O. de Brock and H. Hillege, ADDIS: an automated way to do network meta-analysis

12008-GEM: Hoorn, A.A.J. van, Individualism and the cultural roots of management practices

12009-EEF: Dungey, M., J.P.A.M. Jacobs, J. Tian and S. van Norden, On trend-cycle decomposition and data revision

12010-EEF: Jong-A-Pin, R., J-E. Sturm and J. de Haan, Using real-time data to test for political budget cycles

12011-EEF: Samarina, A., Monetary targeting and financial system characteristics: An empirical analysis

12012-EEF: Alessie, R., V. Angelini and P. van Santen, Pension wealth and household savings in Europe: Evidence from SHARELIFE

13001-EEF: Kuper, G.H. and M. Mulder, Cross-border infrastructure constraints, regulatory measures and economic integration of the Dutch – German gas market

13002-EEF: Klein Goldewijk, G.M. and J.P.A.M. Jacobs, The relation between stature and long bone length in the Roman Empire

13003-EEF: Mulder, M. and L. Schoonbeek, Decomposing changes in competition in the Dutch electricity market through the Residual Supply Index

13004-EEF: Kuper, G.H. and M. Mulder, Cross-border constraints, institutional changes and integration of the Dutch – German gas market

13005-EEF: Wiese, R., Do political or economic factors drive healthcare financing privatisations? Empirical evidence from OECD countries

13006-EEF: Elhorst, J.P., P. Heijnen, A. Samarina and J.P.A.M. Jacobs, State transfers at different moments in time: A spatial probit approach

13007-EEF: Mierau, J.O., The activity and lethality of militant groups: Ideology, capacity, and environment

13008-EEF: Dijkstra, P.T., M.A. Haan and M. Mulder, The effect of industry structure and yardstick design on strategic behavior with yardstick competition: an experimental study

13009-GEM: Hoorn, A.A.J. van, Values of financial services professionals and the global financial crisis as a crisis of ethics

13010-EEF: Boonman, T.M., Sovereign defaults, business cycles and economic growth in Latin America, 1870-2012

13011-EEF: He, X., J.P.A.M Jacobs, G.H. Kuper and J.E. Ligthart, On the impact of the global financial crisis on the euro area

13012-GEM: Hoorn, A.A.J. van, Generational shifts in managerial values and the coming of a global business culture

13013-EEF: Samarina, A. and J.E. Sturm, Factors leading to inflation targeting – The impact of adoption

13014-EEF: Allers, M.A. and E. Merkus, Soft budget constraint but no moral hazard? The Dutch local government bailout puzzle

13015-GEM: Hoorn, A.A.J. van, Trust and management: Explaining cross-national differences in work autonomy

13016-EEF: Boonman, T.M., J.P.A.M. Jacobs and G.H. Kuper, Sovereign debt crises in Latin America: A market pressure approach

13017-GEM: Oosterhaven, J., M.C. Bouwmeester and M. Nozaki, The impact of production and infrastructure shocks: A non-linear input-output programming approach, tested on an hypothetical economy

13018-EEF: Cavapozzi, D., W. Han and R. Miniaci, Alternative weighting structures for multidimensional poverty assessment

14001-OPERA: Germs, R. and N.D. van Foreest, Optimal control of production-inventory systems with constant and compound poisson demand

14002-EEF: Bao, T. and J. Duffy, Adaptive vs. eductive learning: Theory and evidence

14003-OPERA: Syntetos, A.A. and R.H. Teunter, On the calculation of safety stocks

14004-EEF: Bouwmeester, M.C., J. Oosterhaven and J.M. Rueda-Cantuche, Measuring the EU value added embodied in EU foreign exports by consolidating 27 national supply and use tables for 2000-2007

14005-OPERA: Prak, D.R.J., R.H. Teunter and J. Riezebos, Periodic review and continuous ordering

14006-EEF: Reijnders, L.S.M., The college gender gap reversal: Insights from a life-cycle perspective

14007-EEF: Reijnders, L.S.M., Child care subsidies with endogenous education and fertility

14008-EEF: Otter, P.W., J.P.A.M. Jacobs and A.H.J. den Reijer, A criterion for the number of factors in a data-rich environment

14009-EEF: Mierau, J.O. and E. Suari Andreu, Fiscal rules and government size in the European Union

14010-EEF: Dijkstra, P.T., M.A. Haan and M. Mulder, Industry structure and collusion with uniform yardstick competition: theory and experiments

14011-EEF: Huizingh, E. and M. Mulder, Effectiveness of regulatory interventions on firm behavior: a randomized field experiment with e-commerce firms

14012-GEM: Bressand, A., Proving the old spell wrong: New African hydrocarbon producers and the 'resource curse'

14013-EEF: Dijkstra P.T., Price leadership and unequal market sharing: Collusion in experimental markets

14014-EEF: Angelini, V., M. Bertoni, and L. Corazzini, Unpacking the determinants of life satisfaction: A survey experiment

14015-EEF: Heijdra, B.J., J.O. Mierau, and T. Trimborn, Stimulating annuity markets

14016-GEM: Bezemer, D., M. Grydaki, and L. Zhang, Is financial development bad for growth?

14017-EEF: De Cao, E. and C. Lutz, Sensitive survey questions: measuring attitudes regarding female circumcision through a list experiment

14018-EEF: De Cao, E., The height production function from birth to maturity

14019-EEF: Allers, M.A. and J.B. Geertsema, The effects of local government amalgamation on public spending and service levels. Evidence from 15 years of municipal boundary reform

14020-EEF: Kuper, G.H. and J.H. Veurink, Central bank independence and political pressure in the Greenspan era

14021-GEM: Samarina, A. and D. Bezemer, Do Capital Flows Change Domestic Credit Allocation?

14022-EEF: Soetevent, A.R. and L. Zhou, Loss Modification Incentives for Insurers Under ExpectedUtility and Loss Aversion

14023-EEF: Allers, M.A. and W. Vermeulen, Fiscal Equalization, Capitalization and the Flypaper Effect.

14024-GEM: Hoorn, A.A.J. van, Trust, Workplace Organization, and Comparative Economic Development.

14025-GEM: Bezemer, D., and L. Zhang, From Boom to Bust in de Credit Cycle: The Role of Mortgage Credit.

14026-GEM: Zhang, L., and D. Bezemer, How the Credit Cycle Affects Growth: The Role of Bank Balance Sheets.

14027-EEF: Bružikas, T., and A.R. Soetevent, Detailed Data and Changes in Market Structure: The Move to Unmanned Gasoline Service Stations.

14028-EEF: Bouwmeester, M.C., and B. Scholtens, Cross-border Spillovers from European Gas Infrastructure Investments.

14029-EEF: Lestano, and G.H. Kuper, Correlation Dynamics in East Asian Financial Markets.

14030-GEM: Bezemer, D.J., and M. Grydaki, Nonfinancial Sectors Debt and the U.S. Great Moderation.

14031-EEF: Hermes, N., and R. Lensink, Financial Liberalization and Capital Flight: Evidence from the African Continent.

14032-OPERA: Blok, C. de, A. Seepma, I. Roukema, D.P. van Donk, B. Keulen, and R. Otte, Digitalisering in Strafrechtketens: Ervaringen in Denemarken, Engeland, Oostenrijk en Estland vanuit een Supply Chain Perspectief.

14033-OPERA: Olde Keizer, M.C.A., and R.H. Teunter, Opportunistic condition-based maintenance and aperiodic inspections for a two-unit series system.

14034-EEF: Kuper, G.H., G. Sierksma, and F.C.R. Spieksma, Using Tennis Rankings to Predict Performance in Upcoming Tournaments

15001-EEF: Bao, T., X. Tian, X. Yu, Dictator Game with Indivisibility of Money

15002-GEM: Chen, Q., E. Dietzenbacher, and B. Los, The Effects of Ageing and Urbanization on China's Future Population and Labor Force

15003-EEF: Allers, M., B. van Ommeren, and B. Geertsema, Does intermunicipal cooperation create inefficiency? A comparison of interest rates paid by intermunicipal organizations, amalgamated municipalities and not recently amalgamated municipalities

15004-EEF: Dijkstra, P.T., M.A. Haan, and M. Mulder, Design of Yardstick Competition and Consumer Prices: Experimental Evidence

15005-EEF: Dijkstra, P.T., Price Leadership and Unequal Market Sharing: Collusion in Experimental Markets

15006-EEF: Anufriev, M., T. Bao, A. Sutin, and J. Tuinstra, Fee Structure, Return Chasing and Mutual Fund Choice: An Experiment

15007-EEF: Lamers, M., Depositor Discipline and Bank Failures in Local Markets During the Financial Crisis

15008-EEF: Oosterhaven, J., On de Doubtful Usability of the Inoperability IO Model

15009-GEM: Zhang, L. and D. Bezemer, A Global House of Debt Effect? Mortgages and Post-Crisis Recessions in Fifty Economies

15010-I&O: Hooghiemstra, R., N. Hermes, L. Oxelheim, and T. Randøy, The Impact of Board Internationalization on Earnings Management

15011-EEF: Haan, M.A., and W.H. Siekman, Winning Back the Unfaithful while Exploiting the Loyal: Retention Offers and Heterogeneous Switching Costs

15012-EEF: Haan, M.A., J.L. Moraga-González, and V. Petrikaite, Price and Match-Value Advertising with Directed Consumer Search

15013-EEF: Wiese, R., and S. Eriksen, Do Healthcare Financing Privatisations Curb Total Healthcare Expenditures? Evidence from OECD Countries

15014-EEF: Siekman, W.H., Directed Consumer Search

15015-GEM: Hoorn, A.A.J. van, Organizational Culture in the Financial Sector: Evidence from a Cross-Industry Analysis of Employee Personal Values and Career Success

15016-EEF: Te Bao, and C. Hommes, When Speculators Meet Constructors: Positive and Negative Feedback in Experimental Housing Markets

15017-EEF: Te Bao, and Xiaohua Yu, Memory and Discounting: Theory and Evidence

15018-EEF: Suari-Andreu, E., The Effect of House Price Changes on Household Saving Behaviour: A Theoretical and Empirical Study of the Dutch Case

15019-EEF: Bijlsma, M., J. Boone, and G. Zwart, Community Rating in Health Insurance: Trade-off between Coverage and Selection

15020-EEF: Mulder, M., and B. Scholtens, A Plant-level Analysis of the Spill-over Effects of the German *Energiewende*

15021-GEM: Samarina, A., L. Zhang, and D. Bezemer, Mortgages and Credit Cycle Divergence in Eurozone Economies

16001-GEM: Hoorn, A. van, How Are Migrant Employees Manages? An Integrated Analysis

16002-EEF: Soetevent, A.R., Te Bao, A.L. Schippers, A Commercial Gift for Charity

16003-GEM: Bouwmeerster, M.C., and J. Oosterhaven, Economic Impacts of Natural Gas Flow Disruptions

16004-MARK: Holtrop, N., J.E. Wieringa, M.J. Gijsenberg, and P. Stern, Competitive Reactions to Personal Selling: The Difference between Strategic and Tactical Actions

16005-EEF: Plantinga, A. and B. Scholtens, The Financial Impact of Divestment from Fossil Fuels

16006-GEM: Hoorn, A. van, Trust and Signals in Workplace Organization: Evidence from Job Autonomy Differentials between Immigrant Groups

16007-EEF: Willems, B. and G. Zwart, Regulatory Holidays and Optimal Network Expansion

16008-GEF: Hoorn, A. van, Reliability and Validity of the Happiness Approach to Measuring Preferences

www.rug.nl/feb